

Discrimination and Variability of UXO from the DARPA Background Clutter Experiment

Erik M. Rosen
Institute for Defense Analyses
1801 North Beauregard Street
Alexandria, VA 22311
(703) 578-2874 (voice)
(703) 578-2877 (fax)
erosen@ida.org

1.0 Introduction

Data collected during the DARPA Background Clutter Experiment allows one to examine sensor responses corresponding to identical targets buried at four different geographical locations. The four sites, two at Fort AP Hill in Virginia and two at Fort Carson in Colorado were surveyed by six contractors using a variety of sensors. We concentrate here on electromagnetic induction (EMI) data collected by Parsons Engineering using a modified Geonics EM61 sensor. At each site, UXO and UXO-like targets were confined to two 5-m by 100-m swaths. In addition, a 4-meter wide lane contained mines and mine-like targets. Except for five sets of registration targets that were buried for navigational purposes, a 100-m by 100-m undisturbed center-square was delineated as the clutter region. One of the sites, Firing Point 20 at A.P. Hill, has been partially excavated and a number of clutter items have been catalogued.

We present, in this paper, an analysis of the target and clutter signatures from the four test sites. Our focus is on target/clutter discrimination and the variability of target features over the four sites. For the EMI data, features are both energy based and image-based, and include peak signal strength, integrated object density, and circularity. The potential for a given feature to be exploited for target/clutter discrimination will be assessed. We will show that for highly cluttered sites it becomes difficult to associate a given sensor response to the object which gives rise to it. To this end, we will discuss a technique of image segmentation known as region-growing that improves ones ability to accurately measure target features.

We will use our full knowledge of the baseline (or ground truth) to characterize both targets and clutter at all four Backgrounds sites. Image processing techniques allow us to separate, or segment, the data images into distinct regions of targets and clutter. Measurement of target and clutter features can then be done in a more accurate, consistent, and physical manner. We will see that the number of false alarms at each site can be dramatically reduced if a set of features is used to characterize the targets and clutter.

This paper represents the first open publication where specifics of the target baseline will be discussed. The official baseline of the Backgrounds program is being released in coincidence with the publication of this paper.

2.0 The DARPA Background Clutter Experiment

The first phase of the DARPA Background Clutter Experiment took place in the fall of 1996 and was simply a data collection effort. The Waterways Experiment Station collected site characterization data at the four Backgrounds sites. This data included meteorological and soil characterizations. The four Backgrounds locations represent three different soil types and a limited variation of man-made contamination. No attempts are made in this paper to correlate target signatures with site characterization data. Six contractors then brought their sensor systems to the four sites to collect raw data over a variety of target types and clutter. The four sensor types used were magnetometer, infrared (IR), electromagnetic induction (EMI) coils, and ground penetrating radar (GPR).

The second phase of the Backgrounds program is data analysis. Several researchers have obtained portions of the raw sensor data and performed analysis with a concentration on improving algorithms designed to discriminate between UXO and clutter. This is the overarching purpose of the program. A better understanding of clutter signatures and target variability is needed to reduce the number of false alarms that are encountered in both UXO remediation efforts and countermine programs. Indeed, detection of metallic UXO to depths of several meters and metallic mines to depths of several inches is not difficult. The difficulty lies in discriminating between targets and man-made metallic scrap. In typical UXO remediation efforts, more than 80% of the objects that are unearthed are categorized as metallic scrap, or clutter. The latest down-looking sensors used in the Army's countermine programs are detecting both metallic and low-metallic mines at approximately the 90% level, but suffer from unacceptably high false alarm rates. To address this problem, DARPA funded the Backgrounds program to advance the state of the art in sensor design and algorithm development for the purposes of increasing detection rates and lowering false alarm rates.

3.0 Site Design

The site design is shown in figure 1. Its dimensions are 100-m in the north-south direction by 125-m in the east-west direction, where north is magnetic north. The five-meter wide blue lane contains buried UXO targets, including 60-mm and 81-mm mortars, and 105-mm and 152-mm projectiles. The depth of the targets varies from 5-cm to 65-cm, where the smaller targets are buried at shallow depths and the bigger targets at deeper depths. Orientation of the UXO is measured in terms of the angles of inclination and declination, where inclination is the angle into the ground and declination is measured in the plane of the earth's surface. The three orientations used were $(0^\circ, 0^\circ)$, $(0^\circ, 90^\circ)$, and $(90^\circ, -)$, where the first number is inclination angle in degrees and the second number is declination angle in degrees. So an orientation of $(0^\circ, 0^\circ)$ means that the ordnance is lying flat in the ground aligned parallel to the earth's magnetic field. An orientation of $(0^\circ, 90^\circ)$ means the ordnance is lying flat perpendicular to the earth's field. And an

orientation of (90°, -) means the ordnance is sticking straight into the ground in which case the angle of declination is meaningless since it amounts to spinning the ordnance along its long axis as it lies perpendicular to the earth's surface. The red lane contains calibration targets including aluminum plates and iron spheres. The measured response over the red lane targets was limited, and so these targets are excluded from the baseline. The yellow lane contains a mix of low-metal and metallic anti-tank (AT) and anti-personnel (AP) mines, including the low-metal AT M-19, the low-metal AP M-14 and PMN, the metallic AT M15, and the metallic AP OZM-3. They are buried at depths of 2-cm to 15-cm. The orange lane contains more of the same targets encountered in the blue lane, except at depths up to 95-cm. Except for five pairs of registration targets that were buried at known locations in the center-square to aid with registration of the data, the 100-m by 100-m center-square is undisturbed soil containing varying amounts of natural and man-made clutter. Each pair of registration targets consisted of an iron sphere and an aluminum plate buried 5-cm deep and separated by 4 meters. This site design was replicated at the four chosen sites, named Firing Point 20 and Firing Point 22 at Fort AP Hill, and Seabee and Turkey Creek at Fort Carson. The site at Firing Point 22 had to be redesigned because of extremely wet conditions in part of the center-square. There, the part of the center-square that was unnavigable was repeated to the east of the orange lane. And so at FP-22, the site extends 145-m in the east direction as compared to 125-m at the other three sites.

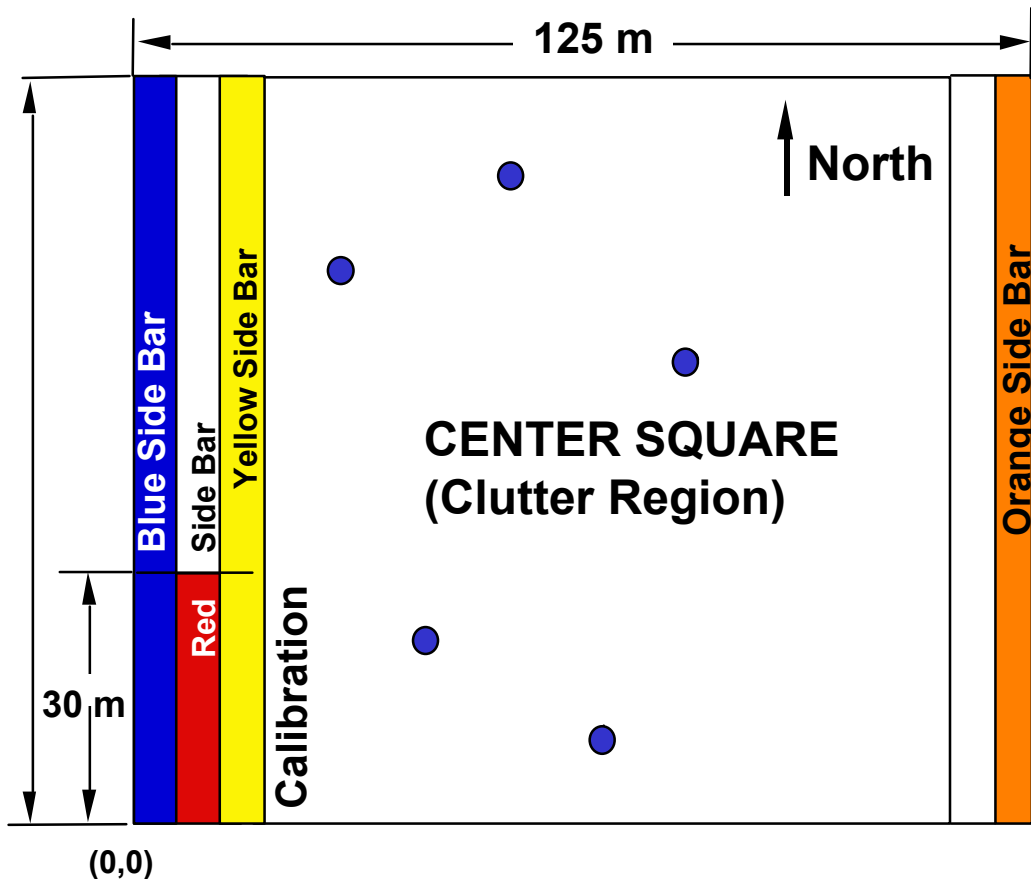


Figure 1: Backgrounds Site Design

4.0 Parsons EMI Data

Our analysis here is on the electromagnetic induction data collected by Parsons Engineering using a modified Geonics EM61 ½ meter coil. The sensor consisted of two ½ meter coils, where one acted as the transmitter, and both acted as receivers. The lower receive coil was located approximately 30 centimeters above the ground, while the top receive coil was located about 70 centimeters above the ground. The integrated time-decay of the induced secondary magnetic field caused by conductors in the ground and received at the top coil is used here in our analysis. Parsons collected data samples using a tick-wheel trigger that amounted to a down-track resolution of approximately 20 centimeters. Cross-track resolution was 50 centimeters. Several sources of location errors were present in the raw data collected at the four sites.

We preprocess the data to account for location errors. Parsons Engineering Science, Inc., used a tick-wheel method of navigation. They traversed the site from 0-m to 100-m north at 1/2-meter intervals in the east-west direction. We use the notation (E0,N0) to indicate a position on the 125-meter by 100-meter site—in this case, east 0 meters, north 0 meters, or the southwest corner of the site. Lining up at the position (E0,N0), Parsons traversed the site, walking to the north until they reached position (E0,N100). At this point, Parsons turned its sensor apparatus around, moved to the east one-half meter, and traversed the site in the southerly direction, from (E0.5,N100) to (E0.5,N0). Parsons repeated this process until most of the site was covered. Cones were placed at the 0-, 50-, and 100-meter points to aid in navigation. We identify and quantify four sources of location error:

- Parallax errors from lining up the sensor head at the beginning and ending of lines (–10 cm to +10 cm in the direction of travel and –5 cm to +5 cm in the cross-track direction)
- Tick-wheel offset caused by the location of the data trigger at the beginning and end of each line (0 cm to +20 cm in the direction of travel for points at the beginning of lines and –20 cm to 0 cm for points at the end of lines)¹
- Meandering from the 1/2-meter east-west lines (–25 cm to +25 cm in the cross-track direction)
- Sensor delay “between the voltage induced in the receiver coil and the associated response measured by the data logger”² (+37.5 cm to +50 cm in the direction of travel)

We reduce the location errors in the direction of travel by aligning sensor maxima with known surveyed locations of targets. An interactive graphical program aids this process. Errors caused by meandering remain in the data. It is the analysis of this modified data set that we present here. Because the Institute for Defense Analyses (IDA) possesses the baseline target information from the DARPA Background Clutter Data Collection Experiment, we were able to perform this location correction. In the case of a blind test, it

¹ This tick-wheel offset results in a location error of –10 cm to +10 cm for points at the center of lines.

² *Background Characterization of the Response of Geophysical Sensors for Subsurface UXO Detection*, Parsons Engineering Science, Inc., January 1997.

is not possible to correct the survey data in this manner; however, an edge filter technique could be used, where misalignment from survey line to survey line is reduced. The technique provides a more accurate anomaly shape but does not result in an improvement in anomaly location relative to baseline location.

We restrict the ground truth to those targets containing a significant amount of metal as to cause a response above the background. We do not expect the EM61 ½ meter coil located 70 centimeters above the ground to detect low-metal mines and so we exclude these targets from the baseline.

5.0 Preprocessing

The first step in our analysis is the formation of data images representing the voltage measured in the upper receive-coil of the modified EM61. We linearly interpolated the data to arrive at a 501 by 251 matrix, where the greater dimension corresponds to the direction of travel (north-south direction). The resolution in the east-west direction is 50-cm and the resolution in the north-south direction is 20 cm. Figure 2 shows the resulting color images of the four sites. The color scale goes from 285-mV to 305-mV for the Seabee and Turkey Creek sites, whereas the color scale goes from 100-mV to 120-mV for the FP 22 and FP 20 sites. We will address this point in the next section.

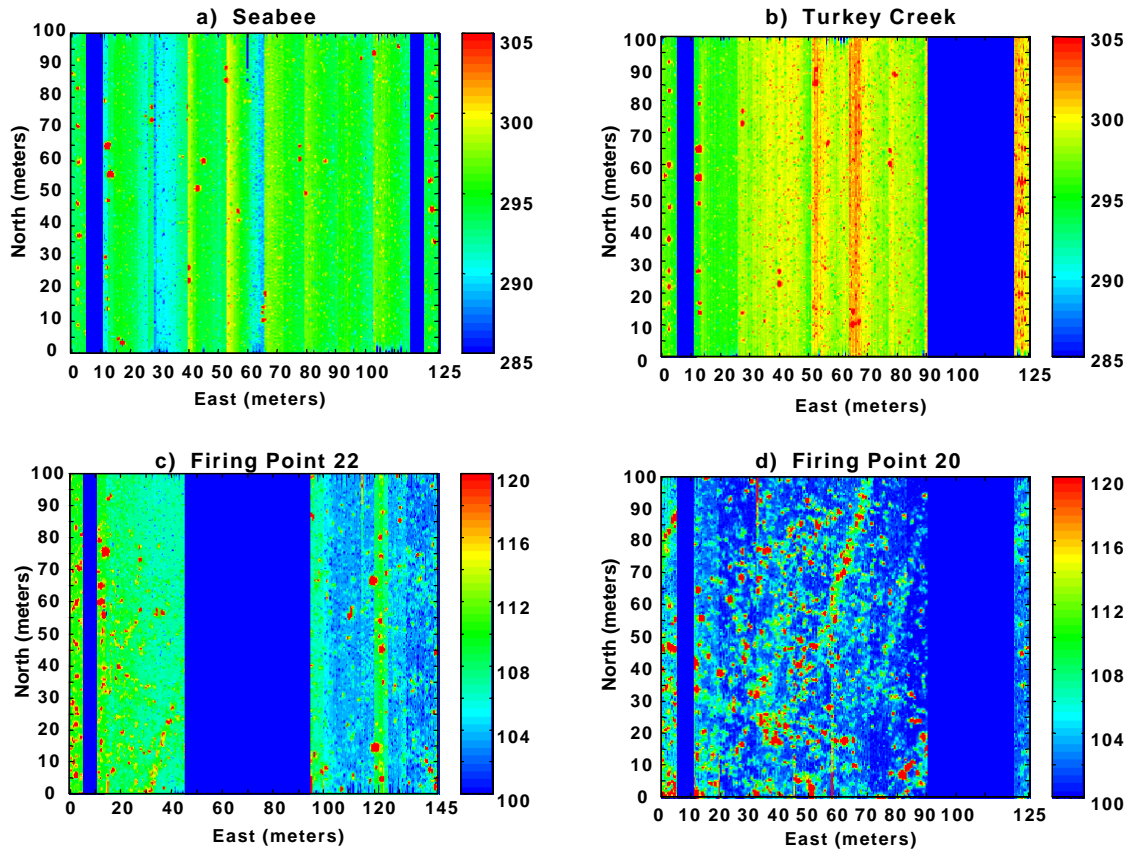


Figure 2: Raw data images

The first thing we observe is the significant but varied response of UXO and mines in the blue, yellow, and orange lanes, as well as the registration targets in the center-square. The responses seem to stand out well against the background, although in some instances they are hard to discern. For instance, in the orange lane (120-m to 125-m east) at FP 20, the target responses are faint. Recall that this is the deep UXO lane.

We also notice that the amount of clutter varies significantly from site to site, with Seabee appearing the least cluttered, and FP 20 appearing the most cluttered. In fact, we will be discussing our observations and results in the context of moving from the least cluttered site (Seabee) to the most cluttered site (FP 20) and so all our 2-by-2 composite figures will be arranged as in figure 2. This will provide consistency and clarity in our presentation of results.

One last observation we make is the presence of vertical streaking patterns at all four sites. These patterns are most noticeable at Seabee and Turkey Creek, and result from discharge of the EM61 battery during the data collection survey. When we look at the distribution of energy at the four sites, this streaking effect has the effect of creating bimodal histograms. This is evident in figure 3, which shows the histograms of the four sites.

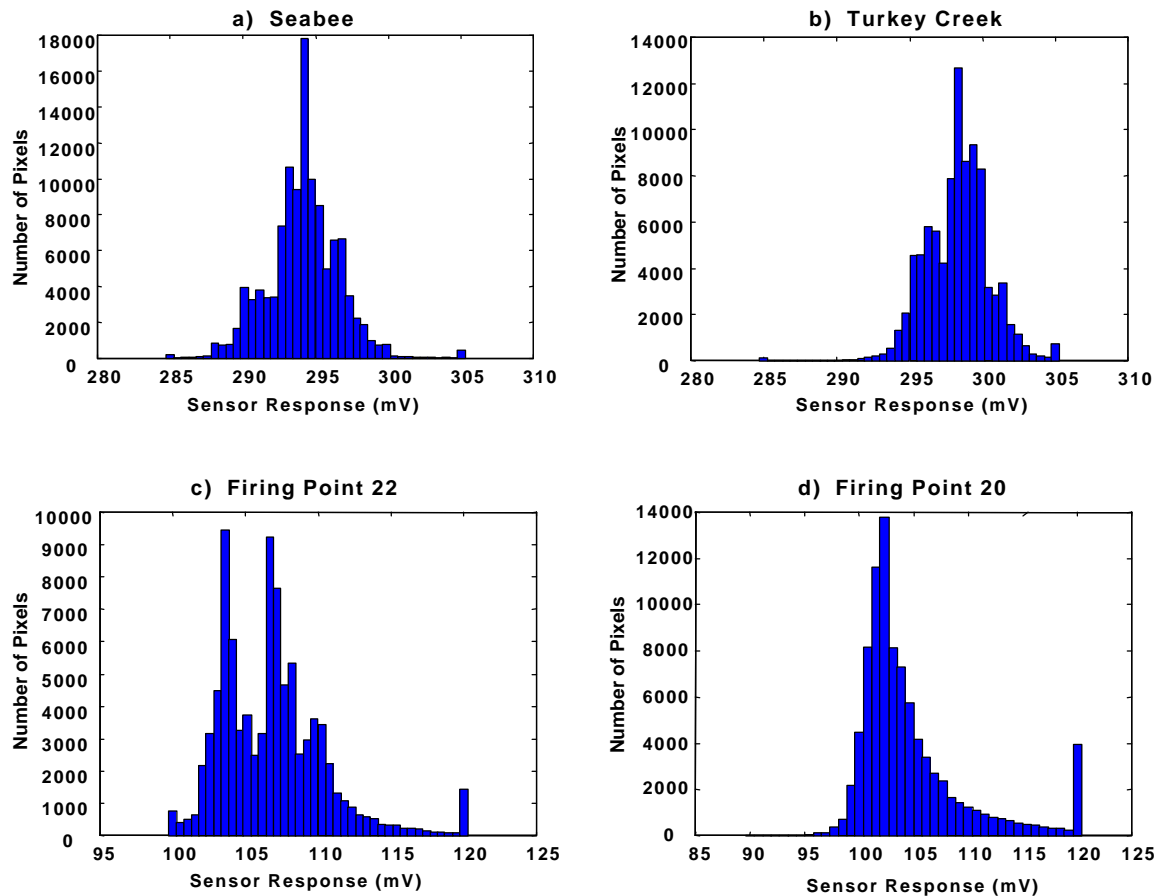


Figure 3: Raw Data Histograms

Notice that both the abscissa and ordinate values vary in each of the site histograms. Our interest is in the shape of the distribution. Referring to the color images of Figure 2, we notice that at FP 22 the background value in the region from 0-m east to 40-m east is consistently higher than the background value in the region from 100-m east to 145-m east. This results in the double peaked histogram of figure 3c. The same phenomena is noticed for the Turkey Creek data, where most background values appear as yellow (from 25-m east to 125-m east), but a number of background values appear green (from 0-m east to 25-m east). This again is reflected in the histogram as shown in figure 3b. For each of the histograms, the bins at the extrema are artificially large due to truncation effects. Thus, we notice at FP 20, there are a significant number of pixels with energy greater than 120-mV, as evidenced by the height of the last bin. This is true to a lesser extent at FP 22. At Seabee and Turkey Creek, there are not a substantial number of pixels with energy values greater than the maximum bin located at 305-mV on the 20-mV dynamic range shown. We compute the mean, standard deviation, minimum, and maximum energy values for the distribution at all four sites. They appear in table 1.

	Mean (mV)	Standard Deviation (mV)	Minimum (mV)	Maximum (mV)
Seabee	294.27	5.87	251.44	836.80
Turkey Creek	298.41	4.69	214.02	742.92
Firing Point 22	107.68	21.16	73.30	1888.10
Firing Point 20	106.10	17.15	83.02	730.39

Table 1: Distribution Statistics for Raw Energy Values

Notice the difference in the mean values between the Fort Carson sites and the Fort AP Hill sites. We noted this difference in the color scales of figure 2. Parsons explain this difference as being attributed to the location of the tracking prism mount. The mount included a large, steel bolt that affected the response at the top coil. The mount was moved in relation to the top coil when the sensor system was transported from Fort AP Hill to Fort Carson. This explains the difference in the background level of the sites.

We preprocess the data to remove the effects of vertical streaking caused by discharging batteries. This also has the effect of changing the mean or background level of all four sites to approximately zero. When we measure anomalies in the data, they will be relative to the same background level. In a given line of data, we eliminate sensor drift (vertical streaking) by subtracting from each value the median for that line. For lines where the majority of pixels have background values, the effect is to zero these pixels. Even for highly cluttered sites such as FP 20, the median pixel value along most lines is equal to the background value, and so for all four sites we should expect the mean to be approximately zero.

Figure 4 shows the data images corresponding to all four sites after performing the transformation just described. Compared to figure 2, we notice that the vertical streaking

is gone, and the anomalies appear against a background value of 0-mV, represented by the color blue. Notice that for all four sub-images, the color scale goes from 0-mV to 10-mV where the units are now relative voltages. There is obviously more clutter present at the Fort AP Hill sites.

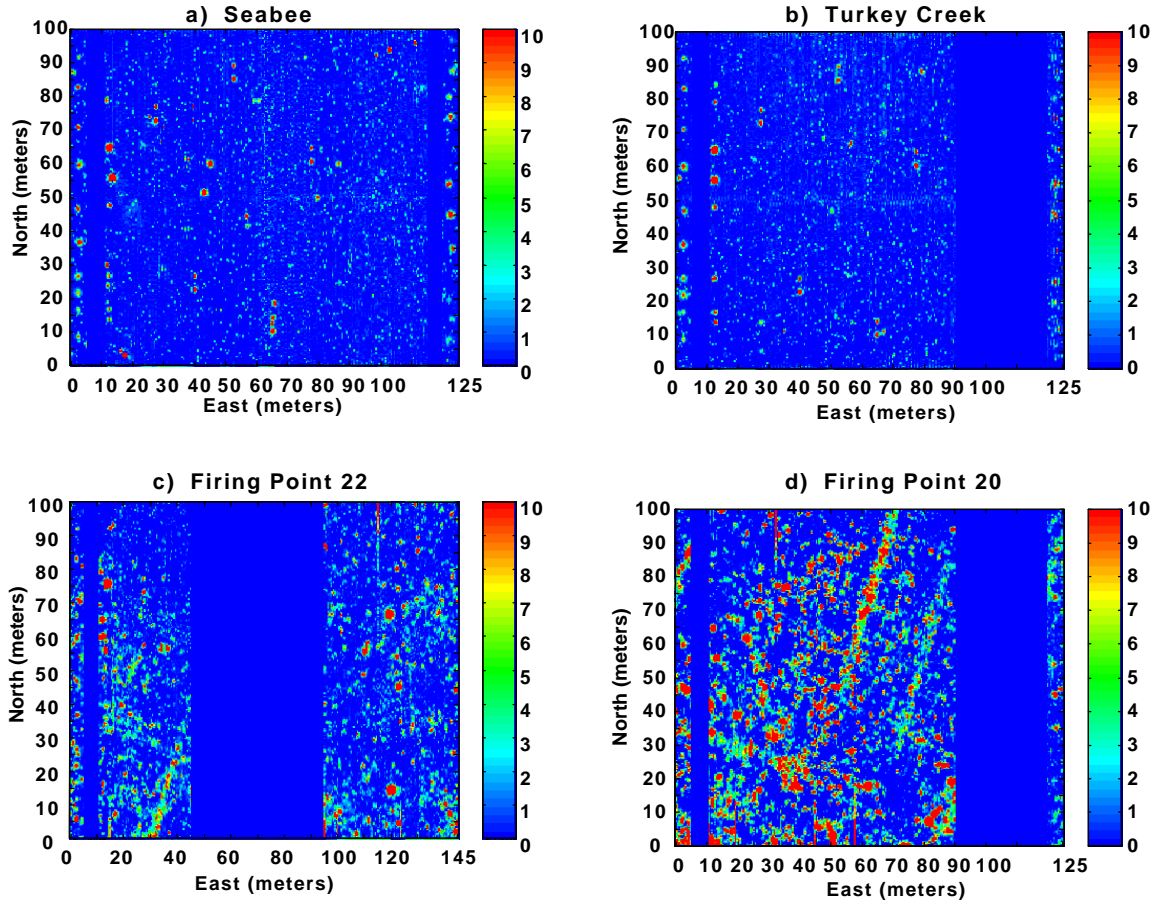


Figure 4: Processed Data Images

When we now examine the histograms and their statistics, we notice that the distributions are much more uniform and centered at 0-mV. Figure 5 shows the site histograms after eliminating the effects of sensor drift.

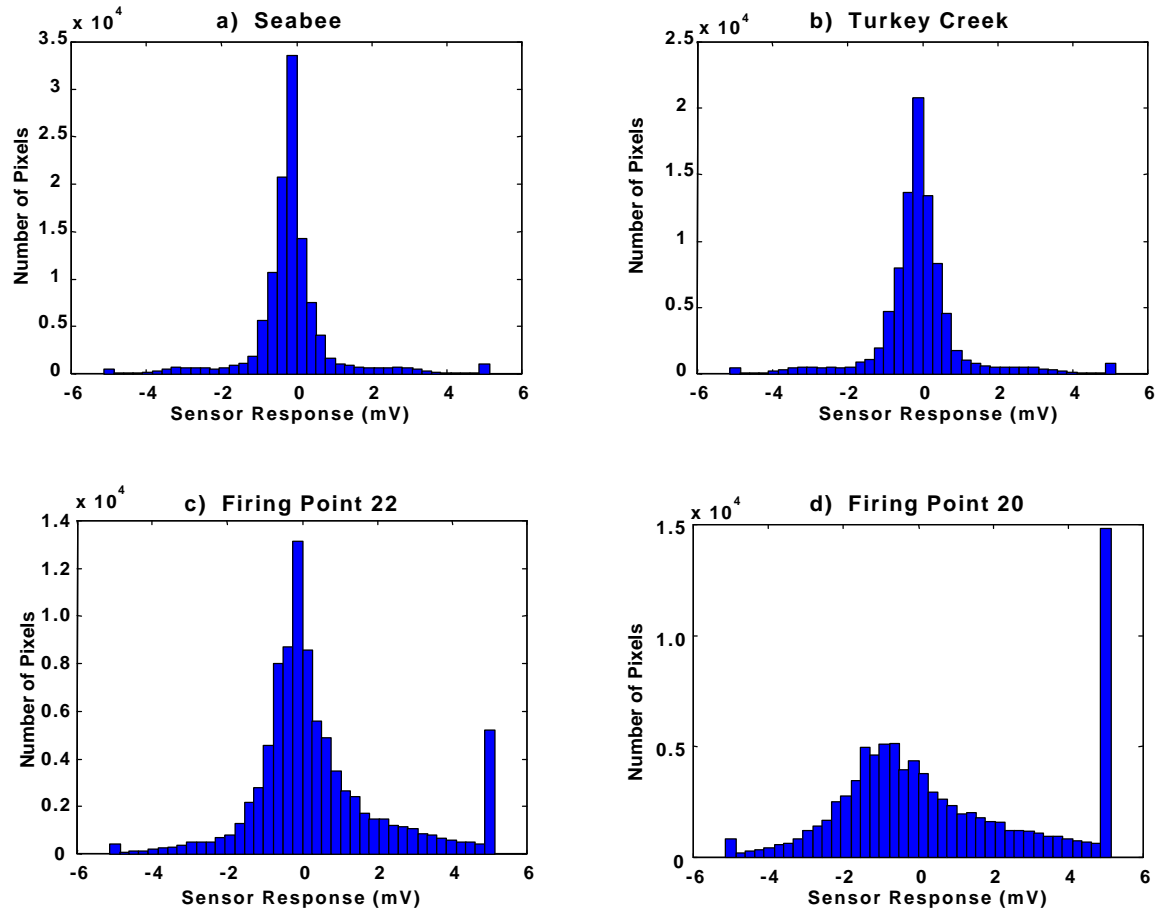


Figure 5: Processed Data Histograms

Notice the distributions are much more uniform now. As with Figure 3, the value of the ordinate varies in each of the histograms. But now the abscissa values are the same. The number of pixel values greater than 5 is significant at FP 20, owing to the apparent large number of clutter items. Referring to table 2, the widths of the distributions for the Fort AP Hill data are significantly greater than the widths of the distributions for the Fort Carson data. This is an indication of the amount of clutter at those sites. Another indication of the amount of clutter present is the mean of the distribution, where the larger the value the more clutter there is. Notice that the maximum signature value at FP 22 is significantly larger than the maximum signature values at the other three sites.

	Mean (mV)	Standard Deviation (mV)	Minimum (mV)	Maximum (mV)
Seabee	0.19	5.51	-43.74	542.53
Turkey Creek	0.12	4.41	-86.20	448.82
Firing Point 22	1.49	20.94	-29.87	1779.50
Firing Point 20	3.00	17.04	-20.43	626.94

Table 2: Distribution Statistics for Processed Energy Values

6.0 Image Segmentation

In order to analyze the data images we must first segment the image into distinct regions of energy. We then attribute these energy 'blobs' as being caused by the presence of some local object buried in the ground. One method of doing this would be to draw a box or circle of some size around each target and associate each of the image pixels that fall within the box or circle with the target. This method of course could not be applied to the clutter sources since they are unknown. In addition, the resultant characteristics of the energy blobs will depend on the size of the box. To this end we use a combination of applied thresholds and mathematical morphology to segment the data images into distinct regions of energy which we then attribute to either a known target or an unknown source of clutter.

The first thing we must do is pick some optimum value of threshold that results in well-separated blobs. This is best illustrated in figure 6 which shows binary images of the Seabee data after applying four different thresholds. For a given threshold, the white pixels correspond to values above the threshold, and the black pixels correspond to values below the threshold.

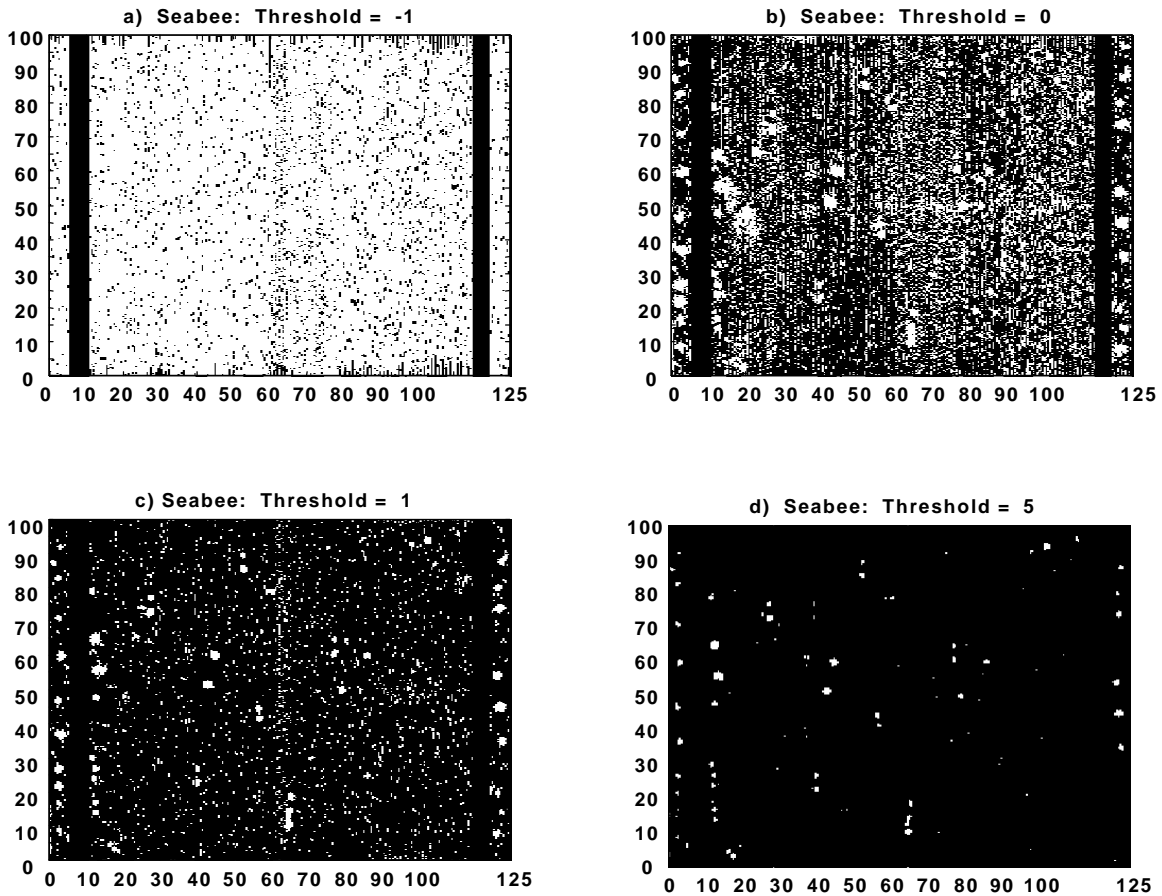


Figure 6: Binary Images at Seabee Site for Varying Thresholds

We choose threshold values that straddle the mean value of the distribution at Seabee, which is 0.19-mV. Notice that for a threshold value of -1-mV, almost all the pixels are white and we are unable to separate the image into distinct regions of energy. At the other extreme when the threshold value is 5-mV it is easy to pick out the energy blobs, but at this point too much information is lost. That is, there are very few pixels on target; or clutter as the case may be. The optimum threshold value lies somewhere in between 0-mV and 1-mV.

We use mathematical morphology, where object shape is the basis for operations on binary images, to define a metric that will aid us in the decision for optimum threshold value. We use the term object to mean connected neighborhoods of pixels. We use a 4-connected neighborhood where pixels connected in the vertical and horizontal directions define an object. This results in more objects of smaller size when compared to an 8-connected neighborhood where diagonally connected pixels are also included in object definition. Referring to figure 6b, we threshold the image at 0 mV and count the number of objects in the image. We repeat this measurement as we increase the threshold value. The result at all four sites is shown in figure 7.

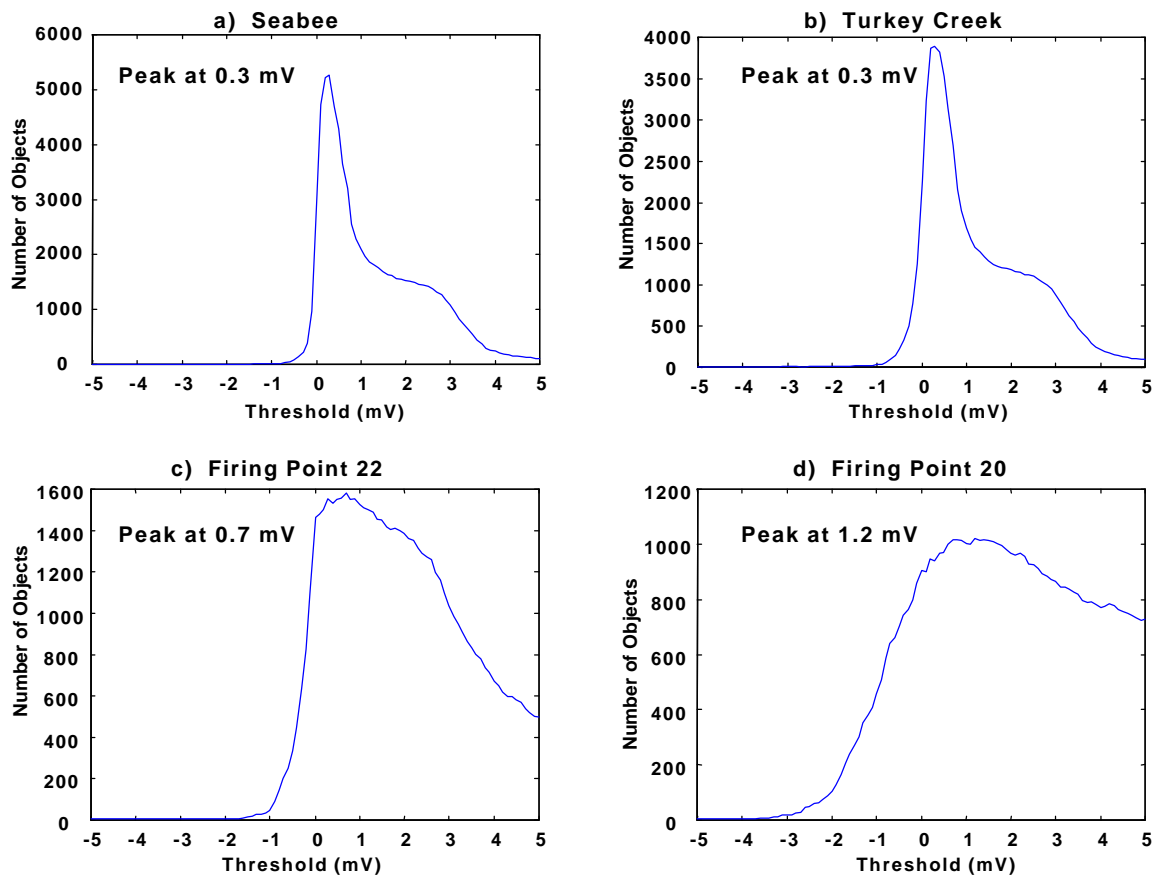


Figure 7: Number of Objects versus Threshold

The threshold value at which the maximum number of objects occurs is indicated in the figure. Notice that the shapes of the curves at Fort Carson are similar with a sharp peak

centered at 0.3-mV, a plateau region extending from about 1-mV to 3-mV, followed by another steep drop in the number of objects. At a threshold of 5-mV, there are very few objects remaining at the Seabee and Turkey Creek sites. Notice that the ordinate is now number of objects and that the range varies from site to site. The peaks of the curves for the data at the Fort AP Hill sites are less sharp, particularly at FP 20. The maximums occur at higher values of threshold when compared to the Fort Carson data. The number of objects remaining when the threshold is set at 5-mV is significant. We now have values of threshold that maximize the number objects in the image and the number of pixels on target.

7.0 Measures of Performance

Since we know the baseline we can compute the probability of detection as a function of threshold. Each target location is mapped to a single pixel of the data image. This pixel has some value of energy associated with it. If at a given threshold, a target pixel value exceeds the threshold, we say that target is detected. Figure 8 shows the results at the four sites.

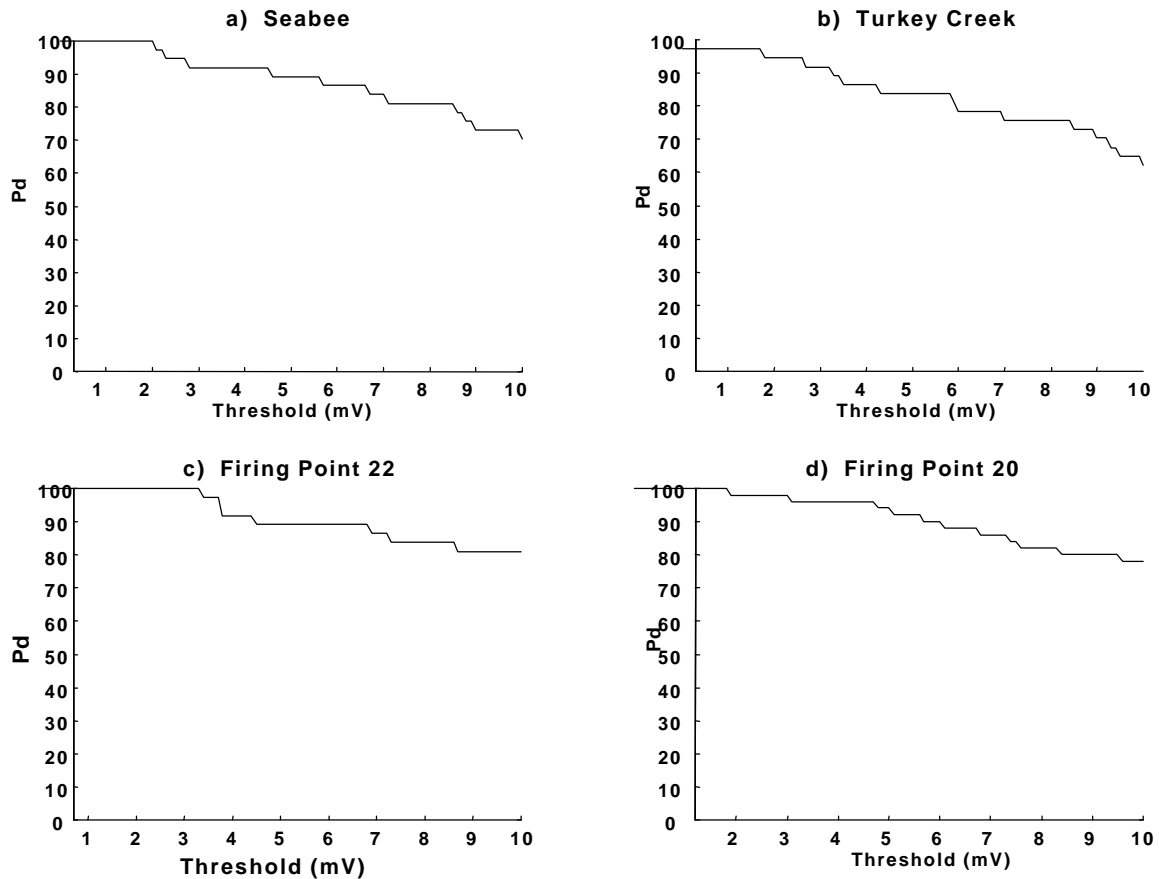


Figure 8: Probability of Detection versus Threshold

We notice that at three of the four sites, the probability of detection (P_d) is 100% at a threshold value of 1-mV. At Turkey Creek, the P_d is about 98%. As the threshold is increased to 10-mV, we notice that targets begin to go undetected, where the steepest drop-off seems to be at Turkey Creek.

Figure 9 shows the probability of detection versus the probability of false alarm, where we define a surrogate for the true probability of false alarm as

$$P_{fa} = N_{fa}/[(A_S - A_T)/\langle A_{fa} \rangle], \quad (1)$$

where N_{fa} is the number of objects that are not associated with targets, $\langle A_{fa} \rangle$ is the average false-alarm area, A_S is the area of the site, and A_T is the area associated with targets. The ratio of the areas is just the opportunities for false alarms. The area of false-alarm pixels is equal to the number of pixels comprising all the clutter objects times the area of a pixel. The site area includes only the parts of the site where Parsons collected data. The target area is equal to the number of pixels on all targets multiplied by the area of a pixel.

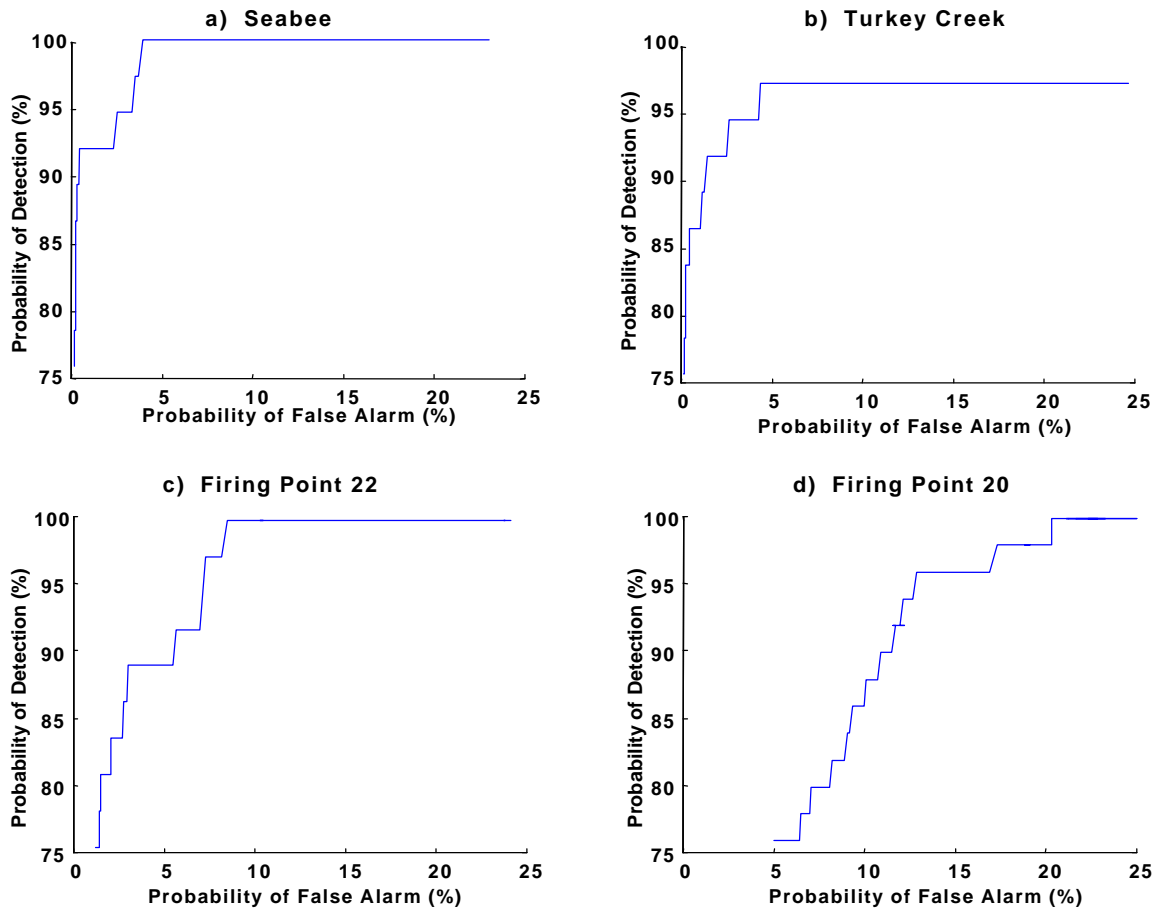


Figure 9: Probability of Detection versus Probability of False Alarm

As we increase the value of threshold and compute P_d and P_{fa} , we move along the curves of figure 9 from the right to the left. We notice that at FP 20, for a given value of P_d , the value of P_{fa} is larger than for the other three sites. As we analyze the data, we attempt to provide a means by which we can, for a given value of P_d , move the value of P_{fa} to the left. We will use our knowledge of the baseline and distributions of object features to show that performance can be enhanced significantly.

As a baseline performance measure, we compute target/clutter distribution functions in the following manner. We threshold the images at the values shown in figure 7. We divide the objects into two classes, namely targets and clutter (by clutter, we mean any object that is not designated as a target). We then count the total number of pixels in each of the two classes that fall in a given energy bin. The result is shown in figure 10 for all four sites. The important point here is that there is complete overlap of target and clutter histograms. In the measurement-space of pixel energy, we cannot discriminate between targets and clutter. Our goal is to find measurement-spaces where separation exists between target and clutter objects.

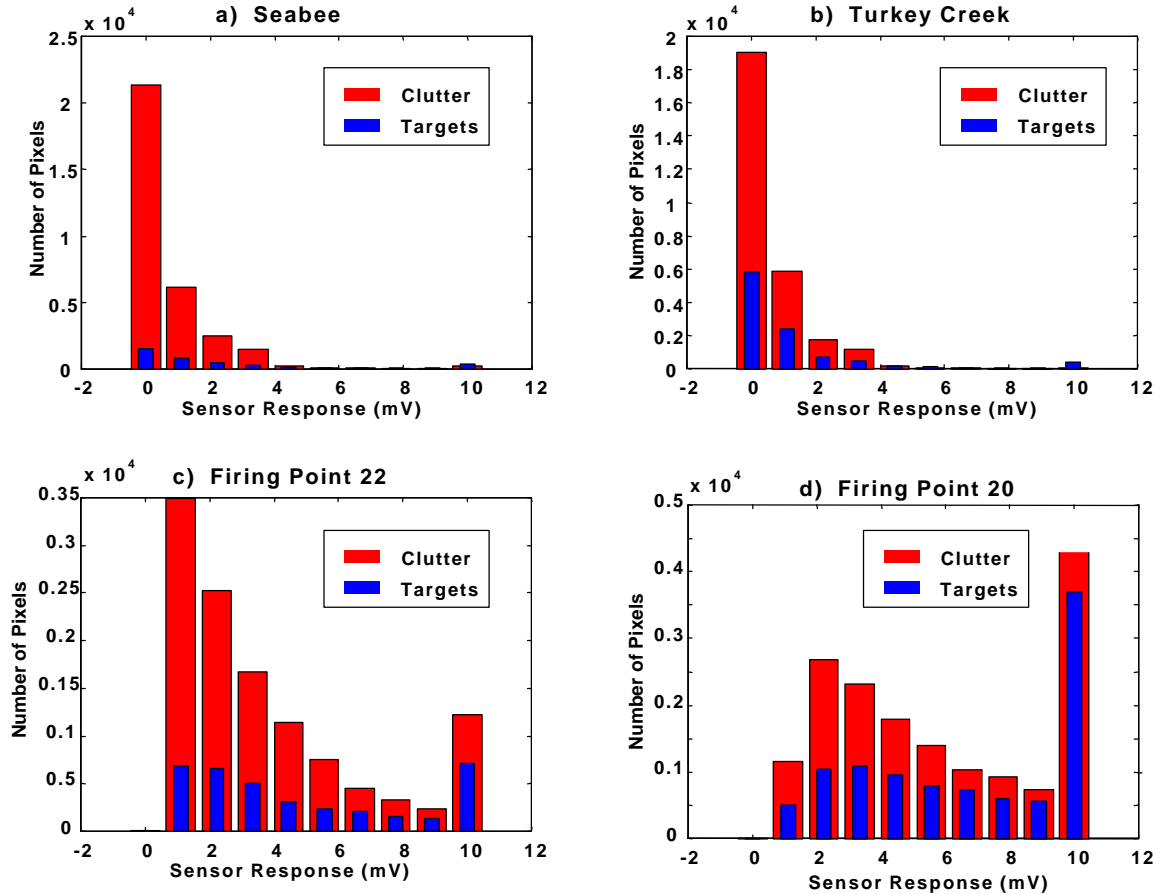


Figure 10: Target/Clutter Energy Distributions

8.0 Image Processing

Having identified individual pixels as belonging to either target or clutter objects, we now examine the results at the four sites. Figure 11 shows target and clutter objects at the four sites after applying the optimum threshold and grouping 4-connected neighborhoods of pixels. The results vary from site to site. At Seabee, almost all targets (shown in white) are easily separated from nearly adjoining clutter objects (shown in red). The exception is two registration targets located at (E65, N12) that appear to be erroneously connected. This problem gets worse at the Turkey Creek site where many of the targets become connected using a threshold of 0.3 mV and the object definition we have chosen. At FP 22 and FP 20, large amounts of clutter, and the presence of additional targets in the center-square exacerbate the problem.

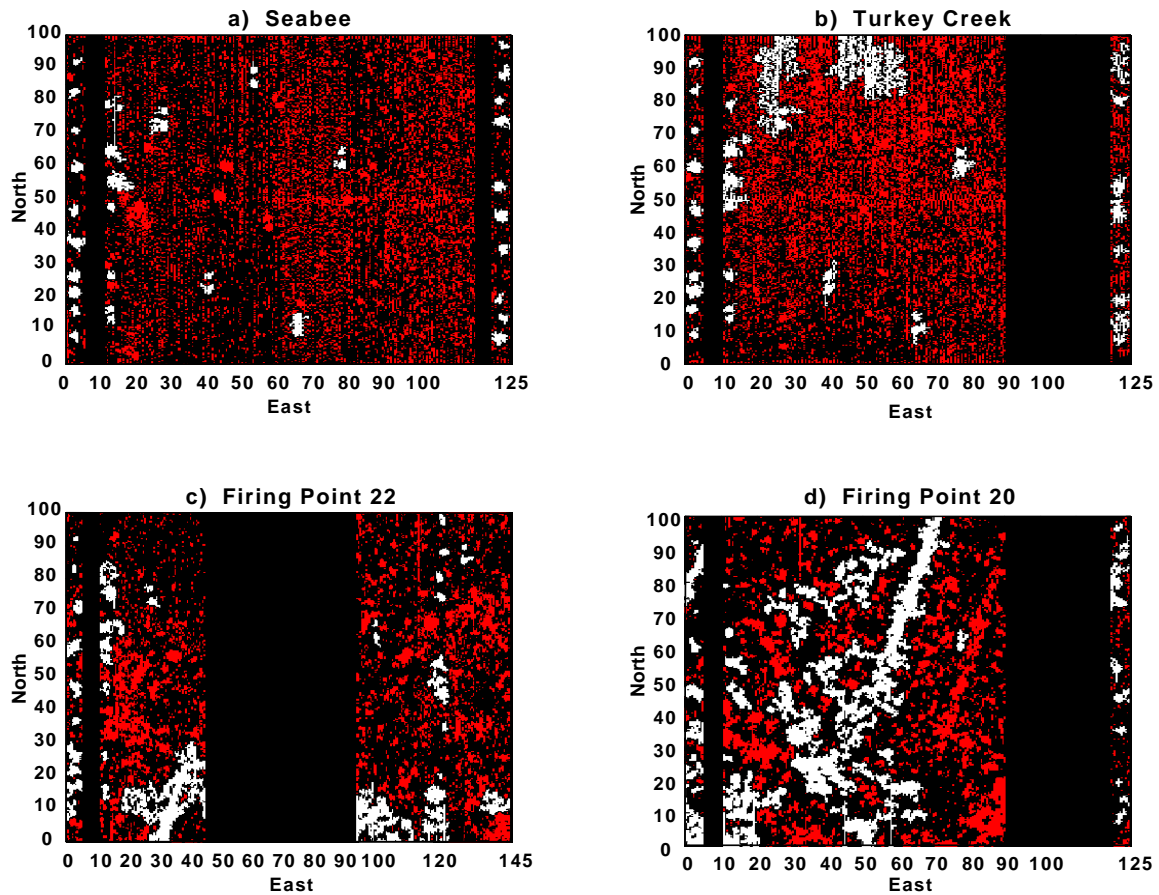


Figure 11: Target and Clutter Objects after Applying Threshold

We will use techniques from mathematical morphology to improve the situation. Namely, we will apply a filter to the binary images of each site that will reduce the number of clutter objects, separate erroneously joined target objects, and smooth the edges of the remaining objects. The detrimental cost of losing detections will be minimal.

We begin by calling pixels with values greater than the optimum threshold ON and pixels with values less than the optimum threshold OFF. In figure 6, the white pixels are ON and the black pixels are OFF. We now apply the following operators to each of the four binary images:

- Erosion
- Thicken
- Threshold.

We observe that objects associated with targets have certain shape characteristics. Namely, target objects have some finite extent as measured by the distribution of pixels comprising them. Target objects also have a spatial distribution that is restricted by the 4-connected neighborhoods under which we define them. Most clutter objects (at sites other than FP 20) have fewer member pixels than target objects and are less compact. To reduce the number of clutter objects in the image, we construct a filter that is simply a three-by-three matrix of ones. We apply the erosion operator with this filter to the four binary images. The erosion operator turns OFF pixels if any of the eight nearest neighbors are also OFF. In this way, we preserve larger, compact target objects and eliminate smaller, irregular clutter objects. For a target or clutter object to pass through the erosion operator, it must have at least nine pixels above the threshold arranged in the same manner as the three-by-three erosion operator. Of course, objects become smaller because of the erosion process. Once the erosion operator is applied, we increase the number of pixels on target by thickening the remaining objects. The thickening operator adds a layer of pixels to the perimeter of remaining objects, without allowing previously unconnected objects to merge together. In this way, we restore target objects to approximately the same size that they were before we applied the erosion operator, without adjoining nearby objects to each other. The thickening operator should prove useful in more cluttered environments where the distance between objects is small. We can apply it any number of times. Successive thickening operations simply increase the size of the remaining objects. For our analysis, we apply the thickening operator twice. Finally, we threshold the image once more, eliminating the pixels added in the thickening process that are below a specified threshold. This threshold could be more or less restrictive than the original threshold. We choose to use the same optimum value of threshold that was used to create the original binary images. Figure 12 shows the results of these image-processing techniques.

Figure 12: Target and Clutter Objects after Morphology

Notice at the Seabee site that the two registration targets are now disconnected. Only one set of connected targets remains at the Turkey Creek site. One of the targets at FP 22 is clearly too large, and several target objects at FP 20 are erroneously connected to neighboring clutter objects. For all four sites, the number of clutter objects is dramatically reduced. The results are summarized in table 3.

	P_d Before	P_d After	# FA Before	# FA After	P_{fa} Before	P_{fa} After
Seabee	100 %	100 %	5219	211	22.9 %	2.8 %
Turkey Creek	97 %	89 %	3859	157	23.9 %	2.5 %
Firing Point 22	100 %	95 %	1543	375	20.3 %	13.7 %
Firing Point 20	100 %	96 %	969	410	22.2 %	21.0 %

Table 3: Performance Measures Before and After Image Processing

Notice that the detection rates before and after image processing are nearly the same, with the worst degradation of performance at Turkey Creek, where three additional UXO in the orange lane are undetected. At FP 22, one of the registration targets is undetected and at FP 20, two targets are not found after performing image processing. The number of false alarms is reduced by a factor of 25 at Seabee and Turkey Creek, by a factor of 4 at FP 22, and by a factor of 2 at Firing Point 20. The probability of false alarm is reduced by an order of magnitude at the Fort Carson sites and by more than 25% at FP 22. At FP 20 however, P_{fa} is not affected greatly by our technique. This is in large part due to the nature of the clutter. At Fort Carson the false alarms are probably caused by fluctuations in the sensor system or in the conductive soil. These false alarms, or clutter blobs, are small in size and have irregular shapes. At FP 20 there are a number of UXO and mine-like clutter items that give rise to energy blobs of substantial spatial extent that cannot be filtered out with the erosion operator.

The next step of our analysis is to measure features of the target and clutter objects. We hope that separation will exist between target and clutter feature distribution functions. Referring to figure 12 though, we expect our feature measurements of some target and clutter objects at FP 22 and FP 20 to be inaccurate due to poor segmentation of the image. In order to improve feature measurements of objects at FP 22 and FP 20, we adopt a more sophisticated method of image segmentation known as region growing by pixel aggregation.

We begin by choosing a threshold value that is located in the upper tail of the energy pixel distributions as shown in figure 5. This high value of threshold ensures that only a few peaks are chosen for the next step of the algorithm. These first few peaks comprise our seed pixels. The regions of energy that are identified by applying this high threshold are now grown until no more pixels can be added. We add neighboring pixels to the peaks if they meet a logical predicate. For our algorithm, we require that neighboring pixels be smaller in magnitude than the seed pixels and that they be greater than some cutoff threshold value. One could allow noise in the data to pass our predicate by introducing a factor less than one by which we multiply the seed pixel before comparing it to neighboring pixels. We choose here to set the factor to one, so that the comparison is exact. This method of segmenting the images assumes that the electromagnetic response is sampled in such a way that the response is singly peaked. Barrow and Khadr showed that the response of an iron sphere as measured by the EM61 1-m coil is double-peaked when the receive coil is 0.15-m from the object, but approaches a single peak at a further distance. Since the resolution here is 50 centimeters in the cross-track direction and 20 centimeters in the down-track direction, and the top receive coil is located 70-cm above the ground; the assumption of this segmentation technique should be sufficient.

The process of adding pixels to the regions continues until no more pixels can be added. At this point, a finite number of regions has been completely defined and are removed from the original image to prevent further classification of them. Next, a lower threshold is chosen which identifies another set of unique peaks. These peaks, or regions, are grown and then removed from the original image. The process continues until all regions above some minimum threshold are uniquely identified. The result of

segmenting the data images for FP 22 and FP 20 using the method of region growing is shown in figure 13b and 13d. We represent targets with white pixels and clutter with red pixels. The results are juxtaposed with the result using thresholds and mathematical morphology as seen in figure 13a and 13c. Notice that by using the method of region growing we more accurately capture the size and shape of the energy anomalies caused by the buried targets. Furthermore, for the FP 22 data, we now detect all 37 targets compared to 35 before. And for FP 20, we detect 49 out of 50 targets compared to 48 before.

Figure 13: Target and Clutter Objects With and Without Region Growing

9.0 Feature Measurement and Target/Clutter Discrimination

Now that we have segmented the image into target and clutter objects, we are ready to measure their features. We use both energy-based features and shape-based features. The energy features are computed using the actual pixel values as measured in relative mV. The shape features do not use the pixel values. They depend only on the spatial distribution of pixels for a given object. Object energy features we measure include

maximum, integrated density, mean, and standard deviation. Object shape features include number of pixels, area, and circularity.

Figures 14 and 15 show the target and clutter object distributions for the integrated density and circularity features. Circularity is computed as the perimeter squared divided by the area. For perfectly circular objects, this translates to a value of circularity equal to 4π , or approximately 12. Notice that the value of the abscissa is the same for each of the sub-images, but that the value of the ordinate varies. In addition, the ordinate is now number of objects, as compared to number of pixels. Notice in figure 14 that the minimum value of the integrated object density for target objects is greater than the IOD value of many clutter objects. The degree to which this is true varies at the four sites. By setting a feature threshold at this minimum value, we can eliminate hundreds of clutter objects, or false alarms. For the feature of circularity shown in figure 15, we notice that the target distributions are centered at 4π , where the peak occurs. We can set feature histogram thresholds at both the minimum and maximum target distribution levels to eliminate more false alarms. When we do this for a set of features, we reduce the number of false alarms at each site significantly.

Figure 14: Target/Clutter Histograms for Integrated Object Density (IOD)

Figure 15: Target/Clutter Histograms for Circularity

Table 4 shows for each feature at all four sites the percentage of false alarms eliminated using our knowledge of the baseline and histogram thresholds.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20
Size (# pixels)	62 %	23 %	66 %	38 %
IOD	84 %	78 %	75 %	33 %
Mean	38 %	66 %	53 %	4 %
σ	46 %	89 %	68 %	10 %
Width	27 %	0 %	50 %	23 %
Length	59 %	23 %	66 %	31 %
Area	49 %	23 %	66 %	32 %
Aspect Ratio	30 %	29 %	13 %	19 %
Perimeter	68 %	23 %	66 %	39 %
Circularity	53 %	23 %	66 %	45 %
Maximum	43 %	90 %	65 %	8 %
Minimum	9 %	1 %	12 %	32 %

Table 4: Percentage of False Alarms Eliminated Using Feature Discrimination

At the Fort Carson sites, the best consistent feature is the integrated object density (IOD). If we simply use the maximum value of the anomalies for discrimination purposes, we would eliminate 90% of the false alarms at Turkey Creek, 43% of the false alarms at Seabee, 65% of the false alarms at FP 22, and only 8% of the false alarms at FP 20. The best discriminators at FP 20 are related to the size of the object. Notice that many of the features in table 4 are related to the spatial size of the object, including size, IOD, and perimeter. We will address feature correlation in the next section.

Table 5 summarizes the effect of using all the features to discriminate between target and clutter objects. We show the number of false alarms before and after feature discrimination. For the Fort Carson sites, the number of false alarms was reduced significantly before discrimination by use of the erosion operator. This is not the case for the Fort AP Hill sites where the image was segmented using region growing techniques.

	# FA Before	# FA After	P _{fa} Before	P _{fa} After
Seabee	211	21	2.8 %	2.7 %
Turkey Creek	157	13	2.5 %	2.4 %
Firing Point 22	2429	383	28.8 %	%
Firing Point 20	1089	435	27.4 %	24.1 %

Table 5: Performance Measures Before and After Feature Discrimination

Notice that the number of false alarms is reduced even further using feature discrimination. An order of magnitude reduction in number of false alarms is achieved at the Fort Carson sites. At FP 22, the number of false alarms is reduced by a factor of 6, while at FP 20 the factor is equal to 2. For all four sites, the probability of false alarm is not substantially reduced. This indicates that most of the false alarms that are being eliminated are small. When we compare table 5 to table 3 we notice that there are more false alarms remaining at the Fort AP Hill sites after feature discrimination than there were after our first attempts at segmenting and filtering the images. The numbers in Table 3 for FP 22 and FP 20; however, are based on an inaccurate segmentation of the image as discussed previously and as illustrated in figure 13.

The final result of eliminating false alarms using target/clutter histogram thresholds is shown in figure 16. There, the remaining false alarms are shown in red, and the targets in white. Compared to figures 12 and 13, the effect of using features to eliminate clutter is noticeable, especially at Seabee and Turkey Creek.

Figure 16: Target/Clutter Objects after Feature Discrimination

10.0 Feature Statistics

Previously, we mentioned that the mean of the target distributions as a function of circularity appear to be centered at 4π . We can quantify this observation by thinking of targets and clutter as two distinct classes and computing statistics of these classes.

We now compute the mean, variance, and variance-normalized distance between class means for each feature. The mean of feature x is defined as

$$\bar{x}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i, \quad (2)$$

where N_j is the number of objects in class j . For objects belonging to different classes, we desire a given feature to take on significantly different values. The larger the difference between the mean of feature x for objects in class j and the mean of feature x for objects in class k , the better the feature is at discriminating between classes j and k . We use the variance of the target and clutter objects as a measure of feature reliability. That is, the

less a particular feature varies within its class, the more reliable it is. The variance of feature x is defined as

$$. (3)$$

We use the variance-normalized distance between class means j and k for feature x to determine the discriminatory capabilities of feature x . It is defined as

$$. (4)$$

Larger values of D_{xjk} indicate a feature that has better discriminatory ability. Remember: for the variances, σ^2 , the units are squared; for the distance, D_{xjk} , there are no units. Tables 6 and 7 show the mean, variance, and d values at all four sites for the IOD and circularity features, respectively.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20
Target Mean	477.8	500.66	595.7	1240.8
Clutter Mean	59.3	26.1	60.0	241.2
Target Variance	827457.0	790539.3	1075696.8	2102963.3
Clutter Variance	75910.4	3070.3	396039.7	399242.9
d	0.44	0.53	0.44	0.64

Table 6: Integrated Object Density (IOD) Target/Clutter Means, Variances, and d metric

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20
Target Mean	13.5	12.5	18.6	18.6
Clutter Mean	7.3	7.6	7.1	12.2
Target Variance	10.1	11.5	69.3	36.9
Clutter Variance	17.3	21.2	33.7	52.3
d	1.18	0.86	1.13	0.68

Table 7: Circularity Target/Clutter Means, Variances, and d metric

With reference to table 6, we notice that target mean is greater than clutter mean at all four sites, and that except for FP 20, the target means are within 20 % of each other. The high target mean value at FP 20 is probably due to the thirteen M15 mines that were found in the center-square. These mines have strong EMI signatures. The target variance is much greater than the clutter variance at all four sites for the IOD feature. The variance is a measure of the reliability of the feature. We would like to identify features that take on similar values for all objects of the same class. We do notice that despite some of the site to site differences in target and clutter means and variances, the distance between the distributions as measured by the value d given in equation 4 is close to 0.5 at all four sites.

Table 7 shows the feature statistics for circularity. At the Fort Carson sites, the target means are close to the value of 4π , while at FP 22 and FP 20 they are equal to 18.6. Remember that we did not apply the erosion and thicken operators for the anomalies at FP 22 and FP 20, and these operators have a tendency to round off the edges of objects. The clutter mean for the circularity feature is consistent at three out of four sites, where at FP 20 it is larger than the other sites. At all four sites, the target mean is greater than the clutter mean. As expected, the target and clutter variance is less at the Fort Carson sites than at the Fort AP Hill sites due to the operator effects mentioned previously. Except for FP 22, the target variances are less than the clutter variances. Finally, the d value is close to one at all sites except FP 20, where it is 0.68. The d value for circularity is twice as great as the d value for IOD at all sites except for FP 20, where the d value is nearly the same.

Features should be uncorrelated with each other. Highly correlated features reflect the same property of the object to be measured and typically add no discriminatory capability. The correlation of features x and y in class j is estimated by

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (5)$$

This quantity is bounded by -1 and +1. A magnitude close to 1 indicates that the features are highly correlated, while a value of 0 indicates the two features are uncorrelated. Table 8 shows the correlation of several feature pairs at the four sites for target objects.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20
Area/Size	0.97	0.96	0.31	0.95
IOD/Size	0.65	0.56	0.50	0.78
IOD/Circularity	0.36	0.15	0.28	0.62
Mean/Circularity	0.24	0.04	0.18	0.33
Maximum/Aspect Ratio	0.18	0.05	0.00	0.13

Table 8: Target Correlation for Feature Pairs

We expect that object size and area would be highly correlated, and indeed table 8 shows that they are, except at FP 22. This result at FP 22 is unusual. The correlation of IOD with size ranges from 0.50 to 0.78. We might expect that the bigger the object in terms of pixels, the bigger the integrated object density. The values of correlation indicate that some objects may be spatially large but have low pixel values compared to objects that are large both spatially and in terms of energy. The last three feature pairs in table 8 have correlation values of less than 0.33. All three pairs consist of one shape-based feature and one energy-based feature. The low correlation values for the last feature pair indicate that if we pair the maximum value of the object with a feature describing the shape of the anomaly we may be able to eliminate two different sets of false alarms and improve our performance.

11.0 Target Variability

We now turn to the variability of specific targets at the four backgrounds sites. Tables 9, 10, and 11 show the maximum target object value at each site for the ordnance in the blue, yellow, and orange lanes. We are interested in target variability across the four sites and do not include a detailed description of each of the targets. The DARPA Backgrounds ground truth has been officially released for those who are interested in more detail than is given here. In the tables that follow, the mean, standard deviation, and normalized standard deviation are computed. The normalized standard deviation is computed as the ratio of the standard deviation and the mean, and is a measure of target variability. The lower this value, the more consistent the target appears at each of the four sites for a given feature. Tables 9, 10, and 11 give results for the feature maximum.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20	Mean	Standard Deviation	Norm of STD
BLU-01	9.93	8.98	7.19	19.38	11.37	5.46	0.48
BLU-02	7.09	6.90	19.70	19.63	13.33	7.32	0.55
BLU-03	8.52	12.54	15.90	19.47	14.11	4.68	0.33
BLU-04	10.92	9.45	13.29	10.03	10.92	1.69	0.15
BLU-05	25.50	31.59	28.49	34.21	29.95	3.78	0.13
BLU-06	16.18	16.44	21.06	47.20	25.22	14.82	0.59
BLU-08	14.08	17.86	21.17	22.76	18.97	3.84	0.20
BLU-09	20.72	16.24	81.66	6.30	31.23	34.16	1.09
BLU-12	16.90	24.42	25.13	30.20	24.16	5.48	0.23
BLU-13	8.96	9.93	8.75	7.35	8.75	1.06	0.12

Table 9: Blue Lane Target Variability for Feature Maximum

Referring to table 9, the value of normalized standard deviation varies from 0.12 for BLU-13 to 1.09 for BLU-09. Notice that two of the targets which vary the least, BLU-13 and BLU-05, have the weakest and strongest peak signal strength. What appears to be an anomalously high maximum for BLU-09 at FP 22 could possibly be explained by examining figure 13b where it is noticed that BLU-09, located at approximately 70-m North is in close proximity to an unknown clutter item. It is possible that the overlap of the two signals contributes to this apparently inconsistent reading.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20	Mean	Standard Deviation	Norm of STD
YEL-05	65.32	-	59.47	63.87	62.89	3.05	0.05
YEL-06	62.21	53.90	72.10	71.72	64.98	8.69	0.13
YEL-09	42.17	42.09	32.65	33.49	37.60	5.24	0.14
YEL-13	43.58	43.45	40.94	44.01	43.00	1.39	0.03
YEL-14	532.69	448.82	528.41	421.15	482.77	56.35	0.12
YEL-16	242.89	231.78	391.84	242.21	277.18	76.71	0.28
YEL-20	45.43	32.25	37.25	15.36	32.57	12.70	0.39

Table 10: Yellow Lane Target Variability for Feature Maximum

Table 10 shows object maximums for the targets in the yellow lane. Recall that these are metallic-cased AT and AP mines buried up to 15 cm deep. Overall, the mines vary much less than the UXO emplaced in the blue lane. The signal strength of YEL-13 is remarkably consistent from site to site. It is interesting to note that the low value for YEL-20 at FP 20 may be related to the size of the identified anomaly as seen in Figure 16. In that figure, YEL-20 appears at 11.5 meters east, 79 meters north, and is much smaller at FP 20 than at the other three sites. It is possible that the mine was not directly passed over and so the strongest signal was never received by the EM61.

	Seabee	Turkey Creek	Firing Point 22	Firing Point 20	Mean	Standard Deviation	Norm of STD
ORA-01	4.56	-	4.55	5.75	4.95	0.69	0.14
ORA -02	4.39	5.93	4.64	5.15	5.03	0.68	0.14
ORA -03	3.56	8.32	3.75	-	5.21	2.70	0.52
ORA -04	18.46	-	25.82	27.36	23.88	4.76	0.20
ORA -05	67.13	47.51	59.53	53.38	56.89	8.41	0.14
ORA -06	19.95	64.39	24.01	23.35	32.93	21.05	0.64
ORA -07	15.99	-	13.72	20.26	16.66	3.32	0.20
ORA -08	8.95	7.08	26.19	10.71	13.23	8.76	0.66
ORA -09	8.62	10.15	14.53	12.38	11.42	2.58	0.23
ORA -10	2.28	-	3.91	6.91	4.37	2.35	0.54

Table 11: Orange Lane Target Variability for Feature Maximum

Target variability as a function of maximum signal of the UXO emplaced in the orange lane shows mixed results. Some of the signals are fairly consistent, while others are not. These results reinforce the notion that even signal strength of identical targets placed at the same depth and same orientation vary widely depending on the conditions of the soil and local geology. Of course sensor stability and other random and systematic errors contribute to the variation in signal strength. The purpose of the Backgrounds program was to minimize these types of errors so that differences in target signatures could be attributed more easily to geological differences at the four sites. Soil characteristics including conductivity and moisture content, as well as the effects of nearby clutter are important factors, as evidenced here.

12.0 Conclusions

We have analyzed one of many sets of data collected during the DARPA Background Clutter Experiment. One of our goals was to show that sensor performance could be improved by considering the features of targets and clutter. With full knowledge of the baseline, we have shown that discrimination of targets and clutter is possible when we consider features of the resulting image objects. Significant reductions in the number of

false alarms can be obtained through the use of image processing and feature discrimination. Furthermore, this reduction in false alarms was performed at no cost to the probability of detection. If we accepted a loss of some detections, we could reduce the number of false alarms even more. The next logical step would be to derive receiver operator curves for various target distributions as a function of one or more features. The entire problem lends itself to a thorough statistical classification analysis.

We have also shown that target variability is significant enough to limit the performance of a particular sensor operating in several different environments. Attempts were made to quantify this limitation by computing variations in target features for each of the emplaced ordnance. We have shown target variation as a function of only one feature here, although the one we chose to show, maximum signal strength, is probably the most readily available.

A similar analysis is on-going for a set of magnetometer data collected at all four sites. A different set of features can be derived from this data, including the depth, size, and orientation of the ordnance. These parameters are computed by modeling the ordnance as a dipole and fitting the target and clutter signatures to the model. These physical features derived from the magnetometer data may be combined with some of the EMI features to further enhance performance.

13.0 References

George, Vivian, and Altshuler, Thomas W., "Design of the DARPA Background Clutter Data Collection Experiment," Proceedings of the UXO Forum, 1997.

George, Vivian, Altshuler, Thomas W. and Rosen, Erik M., "DARPA Background Clutter Data Collection Experiment Excavation Results at Firing Point 20," Proceedings of SPIE, Volume 3392, April 1998.

Rosen, Erik, and Altshuler, Thomas W., "Clutter and Target Signature Statistics from the DARPA Background Clutter Experiment," Proceedings of SPIE, Volume 3392, April 1998.

Serra, Jean, "Introduction to Mathematical Morphology," Computer Vision, Graphics, and Image Processing 35, pp. 283–305, 1986.

Kenneth R. Castleman, Digital Image Processing, Prentice Hall, 1996.

Arthur R. Weeks, Jr., "Fundamentals of Electronic Image Processing", 1996.

T.W Altshuler, A.M. Andrews, and D.A. Sparrow, "Mine and UXO Detection: Measures of Performance and Their Implication in Real-World Scenarios," *Proceedings of SPIE* 3079, p. 284, April 1997.

Bruce Barrow, Nagi Khadr, and Herbert H. Nelson, "Performance of Electromagnetic Induction Sensors for Detecting and Characterizing UXO," *Proceedings UXO Forum 1996*, p. 308, 1996.